ARTIFICIAL INTELLIGENCE FOR NEW DRUG DESIGN

DEEP LEARNING FOR LIGAND-BASED DE NOVO DESIGN IN LEAD OPTIMIZATION: A REAL LIFE CASE STUDY

IKTOS

Iktos facts

Our company

Paris-based AI company founded late 2016

30 people

Specializing in AI applied to chemistry:

- Deep Generative models for *de novo* drug design
- Data-driven retrosynthesis

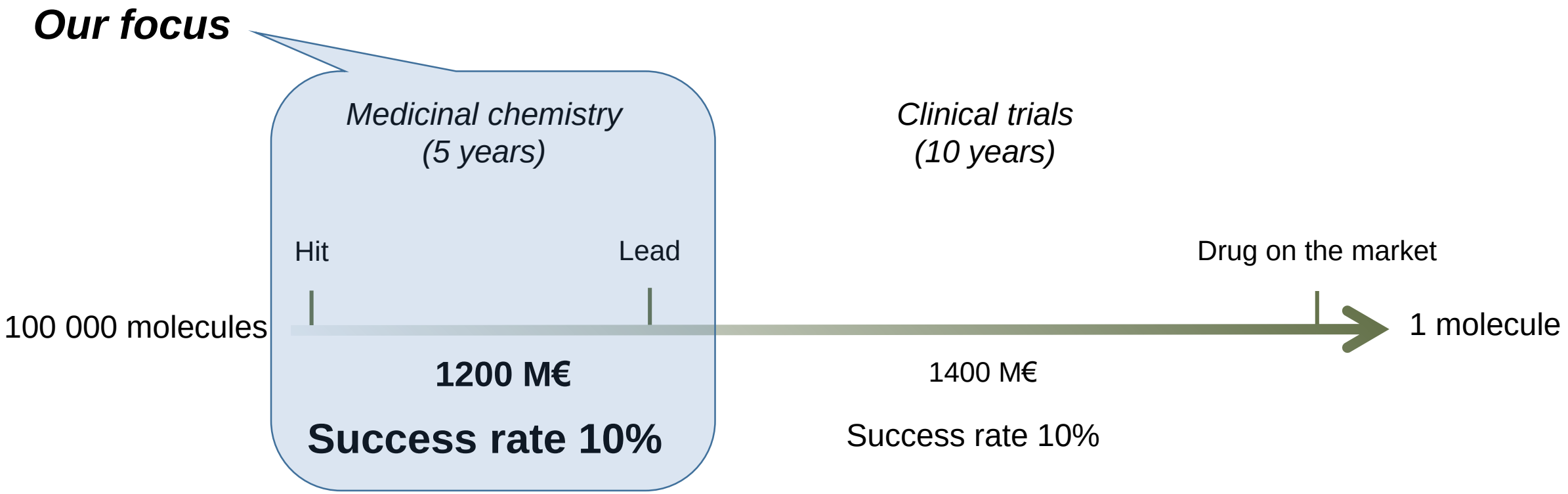
Concrete real-life experience of delivering value to drug discovery programmes: ~15 projects delivered or in progress

Business model: services, research collaborations, software

Our customers

MedChem:
A difficult job!

life cycle: lc



Less than 1% chance to find a new drug…

➔ ***Our objective: use AI to identify the optimal molecule, faster, cheaper, with higher probability of success!***

Solving the Rubik's cube:

- Simultaneous optimization on activity, potency, ADME, tox, selectivity…

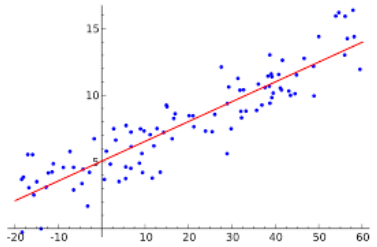
☹ : Gain on one objective usually results in loss on the other ones

- The chemical space is huge (10^{60}). Does the solution even exist? Can we ever find it?

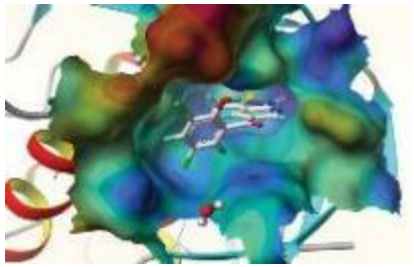
silico approa

Predictive approaches:

- QSAR, data science

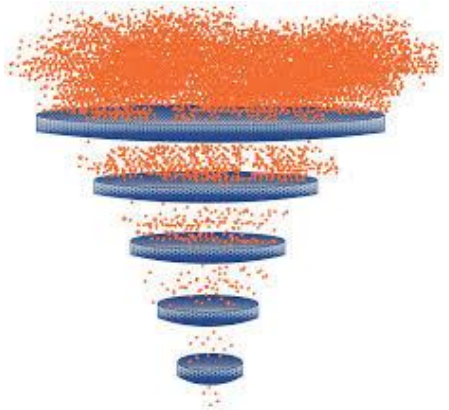


- Molecular modeling



Those technologies can only predict:
They tell you what not to do, but they do not tell you what to do!

Compound *de novo* design approaches:



Virtual screening: **Brute-force**

- Limited by computational power and space
- Only very small portions of the chemical space are explored (10^9 vs 10^{60})
- Virtually Zero chance of finding “the” molecule

Evolutionary algorithms

- Slow, limited diversity, compound feasibility issues

A filtering, rather than designing approach!

Artificial intelligence: changing the rules of the game!

ep Neuronal Net

Automatic colorization of black and white images

Automatic image caption generation

Automatic game playing

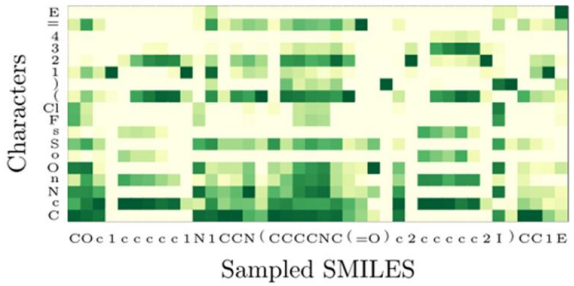
Automatic picture generation

Why not generate molecules instead of images of cats?

Generative

Sequence based

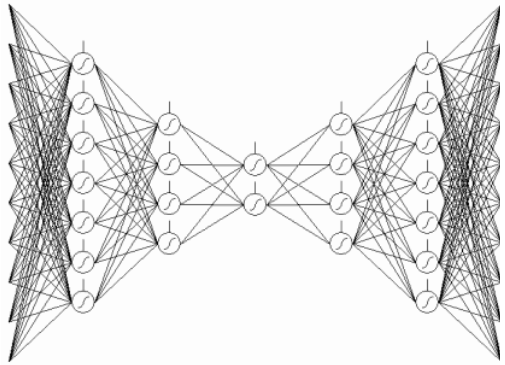
Long Short Term Memory
- LSTM (Characters or Grammar)
- LSTM with Reinforcement learning



AstraZeneca 2017

Latent space

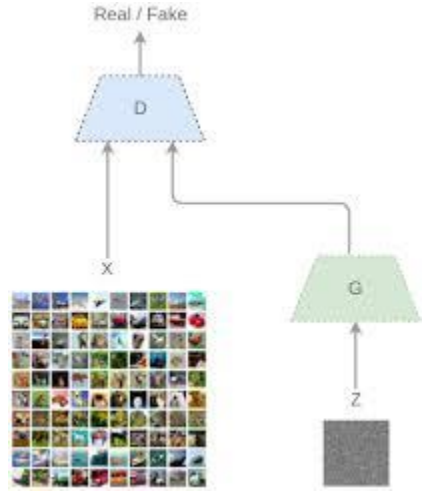
Variational Auto-Encoder
- CVAE, GVAE, SD-VAE



Toronto
Harvard 2016

GAN

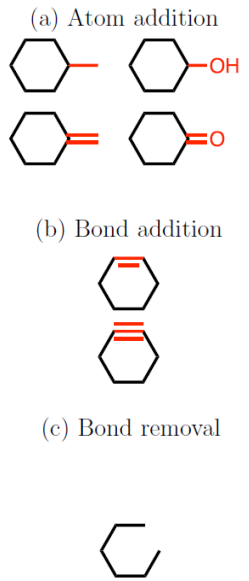
Generative Adversarial Network
- ORGAN



Insilico Medicine 2017

DQN

Deep Q-Networks



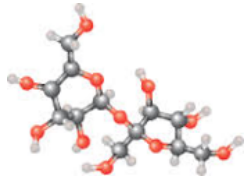
Google 2018

Recent review: Hongming Chen, Ola Engkvist, Yinhai Wang, Marcus Olivecrona and Thomas Blaschke, *The rise of deep learning in drug discovery* Drug Discovery Today, 1-10, February 2018; Youjun Xu, Kangjie Lin, Shiwei Wang, LeiWang, Chenjing Cai, Chen Song, Luhua Lai, Jianfeng Pei, *Deep learning for molecular generation*, FutureMed. Chem. January 2019; Daniel C. Elton, Zois Boukouvalas, Mark D. Fuge, Peter W. Chung, Deep learning for molecular generation and optimization - a review of the state of the art, arXiv March 2019

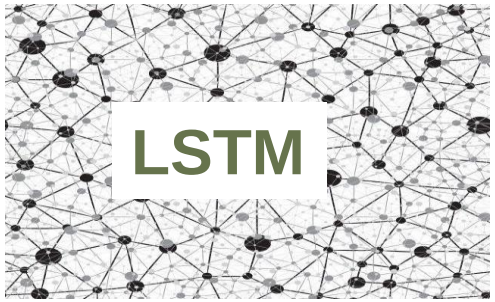
or MPO: LSTM

Generative model (AI)

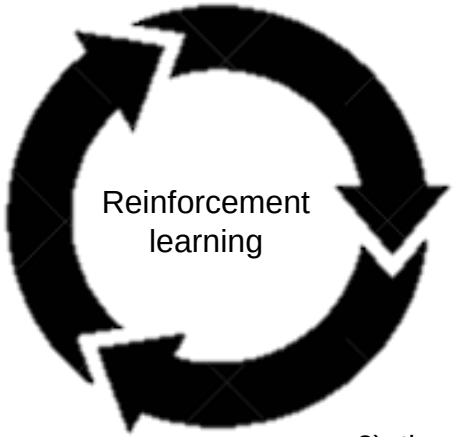
Policy Gradient (AI)



Trained on 86 million of molecules



1) Molecules are generated

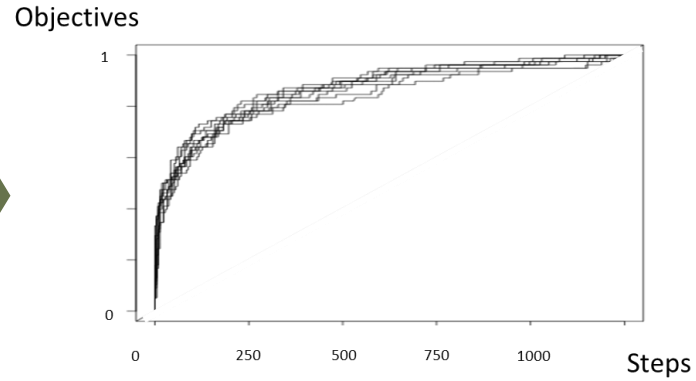


2) molecules are scored by the multi-objective fitness function

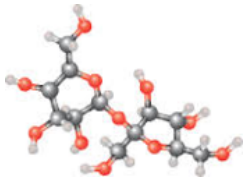
- ✓ QSAR models
- ✓ Docking score
- ✓ Metrics, descriptors
- ✓ Sub-structures

…

Traditional approaches

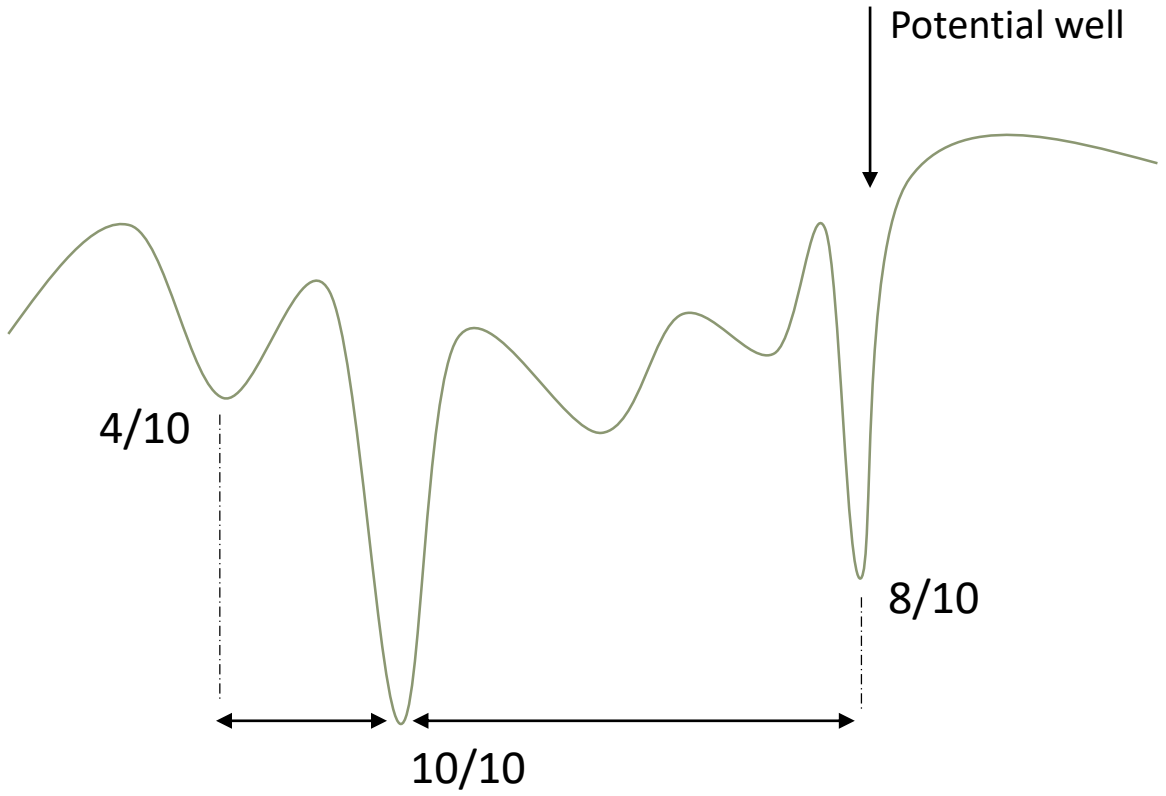


3) the weights of the model are adjusted to maximize the probability of generating molecules similar to those maximizing the global score using a policy gradient algorithm.



Major MPO

10 parameters to optimize for instance



A molecule which fixes 8 parameters over 10, doesn't mean it is closer to the solution than a molecule which fixes only 4 parameters!

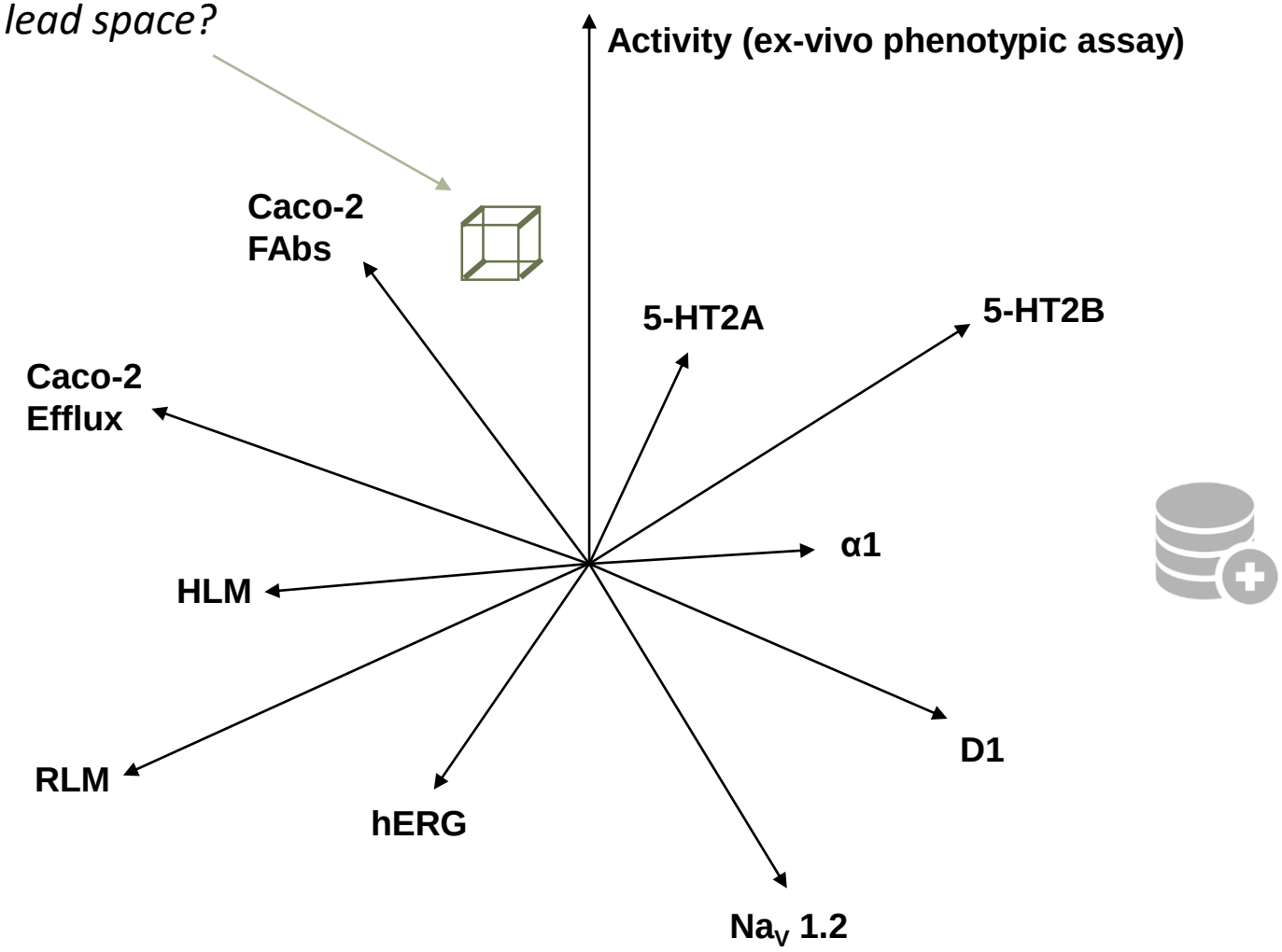
- Valid SMILES rate[1]
- Synthetic access, compound "quality"
- QSAR Applicability Domain
- Diversity / Novelty[2]
- Global fitness function… or not..

1 – Mariya Popova, Olexander Isayev, Alexander Tropsha *Deep Reinforcement Learning for De-Novo Drug Design* arXiv Nov 2017
2 – Kristina Preuer, Philipp Renz, Thomas Unterthiner, Sepp Hochreiter, Günter Klambauer *Fréchet ChemblNet Distance: A metric for generative models for molecules* arXiv March 2018

Real life MPO
Case study

PO Project cl

Where is the Optimal lead space?

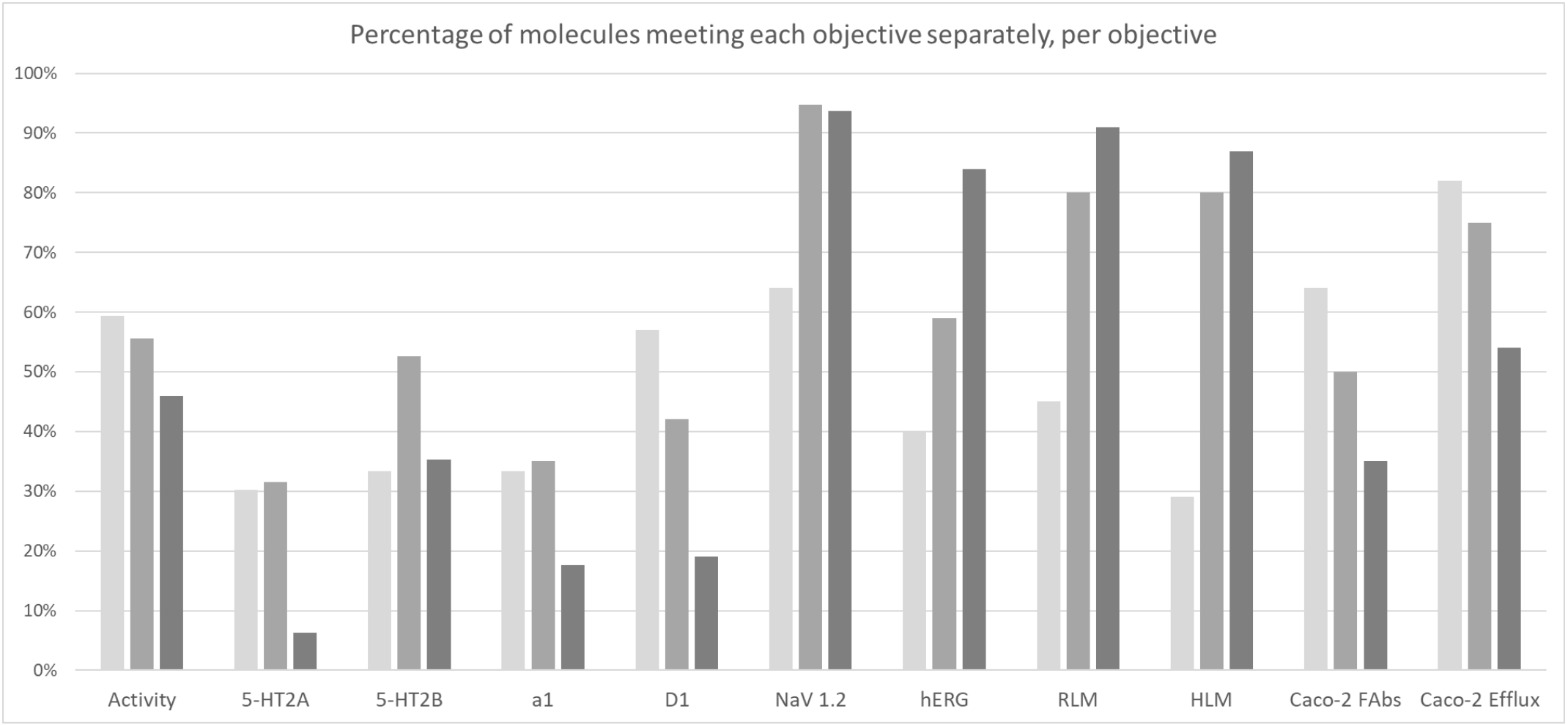


Project facts and figures

- Late-stage Lead Optimization Project
- Undisclosed target, complex, expensive ex-vivo phenotypic assay
- 11 objectives
- 880 molecules (251 measured on Activity)
- Several years of research

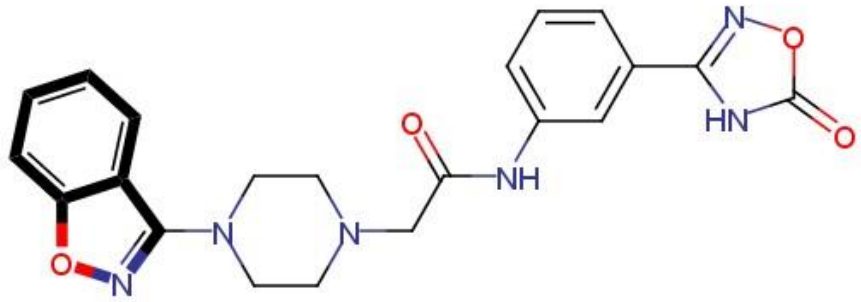
➔ ***No molecule meeting simultaneously the 11 objectives of the blueprint…***

ashboard alo



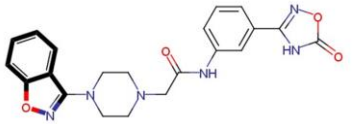
ompound in t

Activity	5-HT2A	5-HT2B	a1	D1	Na_V 1.2	hERG	RLM	HLM	Caco-2 FAbs	Caco-2 Efflux
194.0	20.0	18.0	1.0	4.0	0.0	19.0	82.85	63.35	88.99	26.2



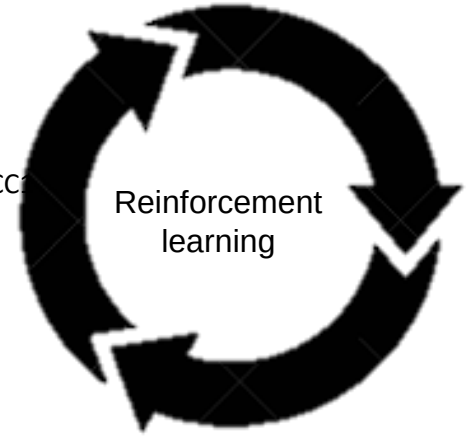
Presence of a 1,2-benzoxazole moiety which appears in 61% of cases in the whole dataset and in 78% of the last 50 molecules

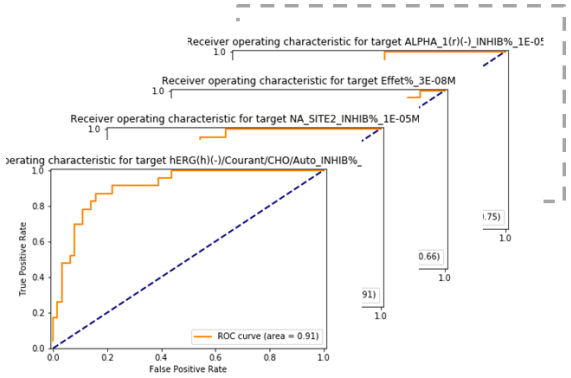
ompounds: t

Activity	5-HT2A	5-HT2B	α1	D1	Na_V 1.2	hERG	RLM	HLM	Caco-2 FAbs	Caco-2 Efflux	
194.0	20.0	18.0	1.0	4.0	0.0	19.0	82.85	63.35	88.99	26.2	 mol 732
83.0	69.0	-25.0	45.0	6.0	13.0	6.4	69.04	31.93	97.6	1.96	mol 663
46.0	46.0	69.0	14.0	14.0	-14.0	25.8	60.28	25.43	98.86	0.75	mol 559
48.0	71.0	48.0	12.0	14.0	39.0	25.0	68.83	33.58	99.37	0.39	mol 555
115.0	76.0	15.0	37.0	-3.0	-13.0	5.4	80.82	83.54	72.24	12.3	mol 550
46.0	6.0	44.0	29.0	-11.0	20.0	12.4	93.11	78.36	73.8	34.1	mol 435

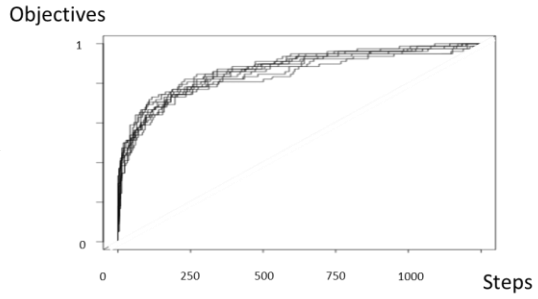
ual compour

Iktos molecules generator





11 Independent QSAR models

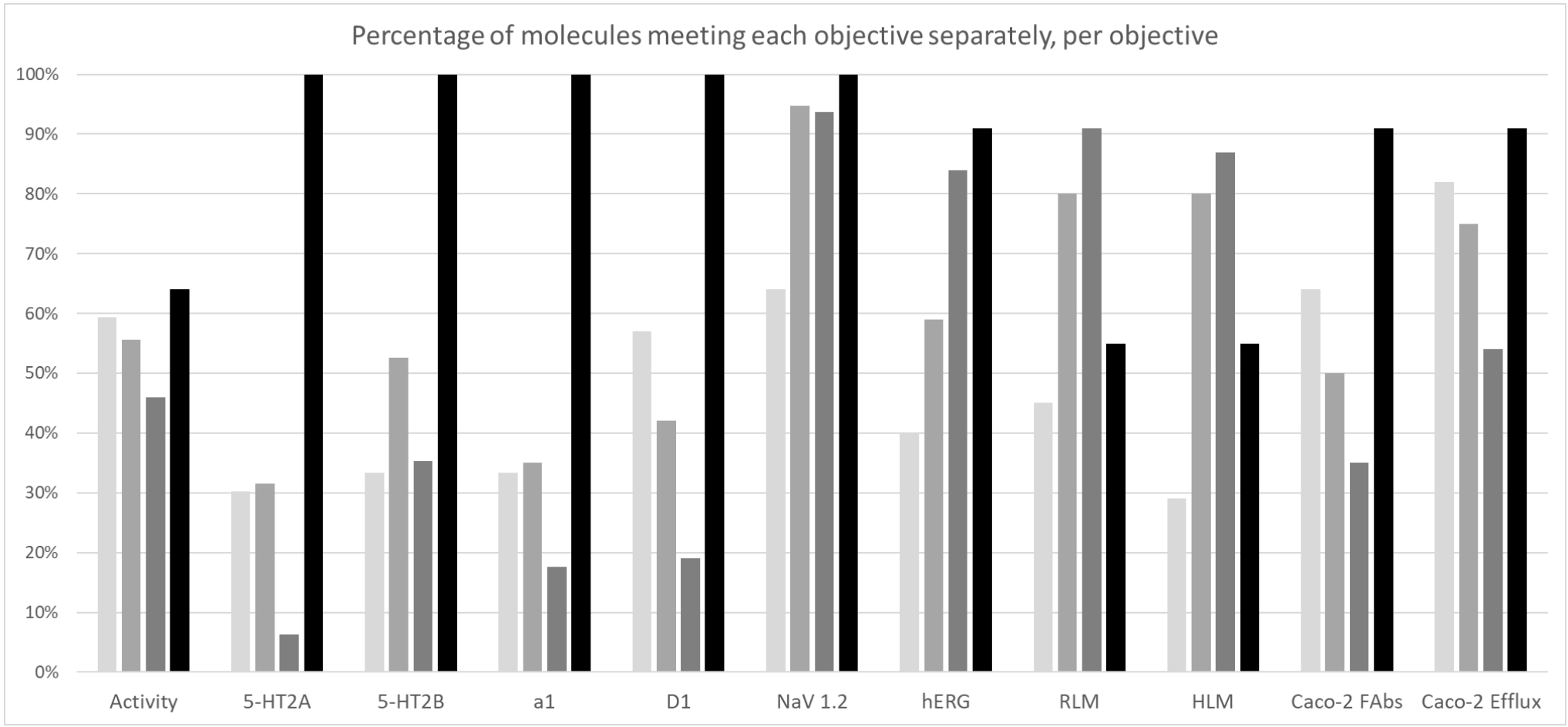


150 virtual “ideal” compounds were designed

Predicted Activity	Predicted 5-HT2A	Predicted 5-HT2B	Predicted α1	Predicted D1	Predicted Na_V 1.2	Predicted hERG	Predicted RLM	Predicted HLM	Predicted Caco-2 FAbs	Predicted Caco-2 Efflux

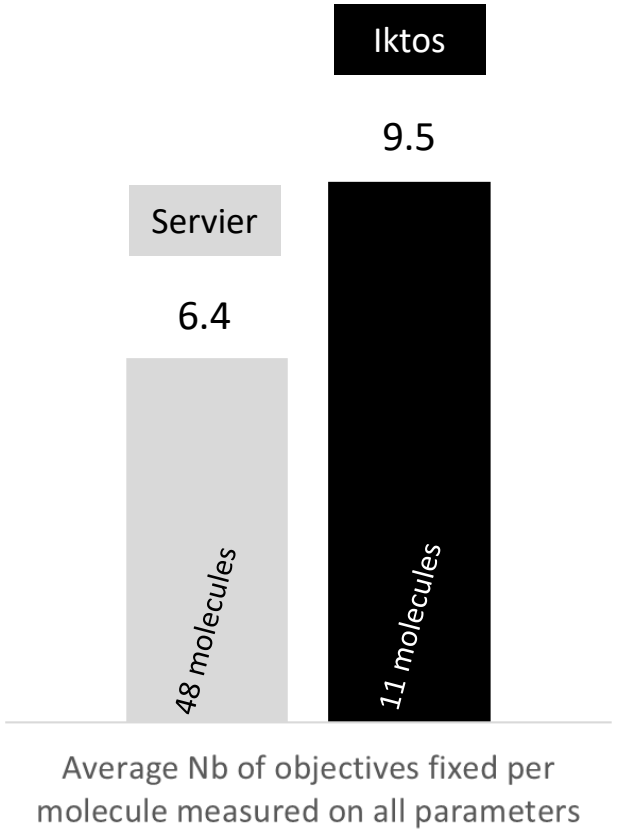
11 compounds were synthesized and tested

tal results af



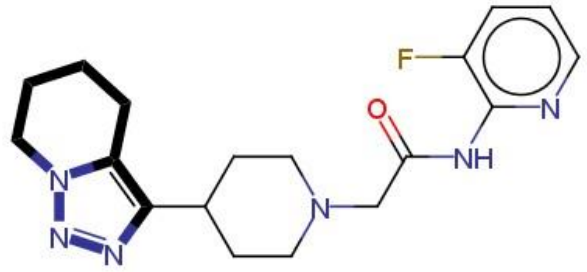
For the majority of objectives, Iktos molecules have a much higher "in blueprint" rate

MPO sur



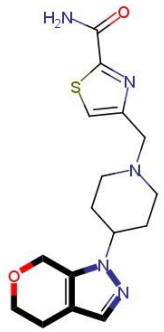
Best AI designed compound

Activity	5-HT2A	5-HT2B	a1	D1	Na_V 1.2	hERG	RLM	HLM	Caco-2 FAbs	Caco-2 Efflux
83	7	18	7	-9	2	3	57	75	97	7

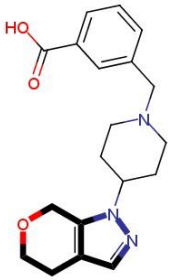


Presence of a [1,2,3]triazolo[1,5-a]pyridine moiety which appears only 6 times in the initial dataset, all having issues on CACO-2 (Permeability & Efflux)

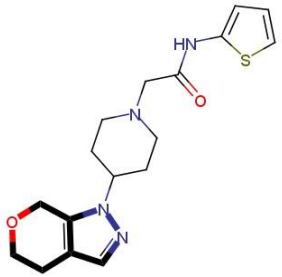
AI designed



Active 9/11

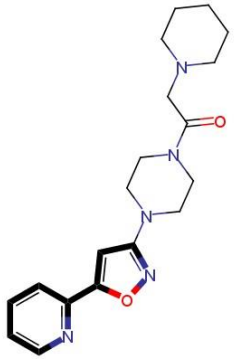


Active 9/11



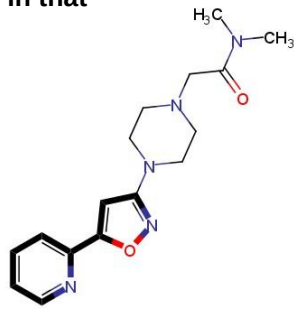
Active 10/11

Presence of a pyrano[3,4-c]pyrazole moiety which appears 5 times in the initial dataset



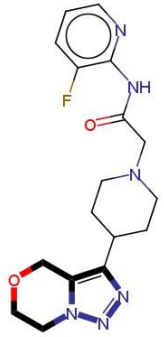
Borderline* 10/11

First introduction of an aliphatic group in that position

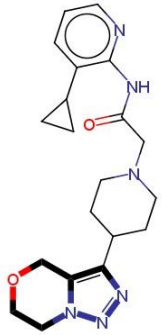


Inactive 8/11

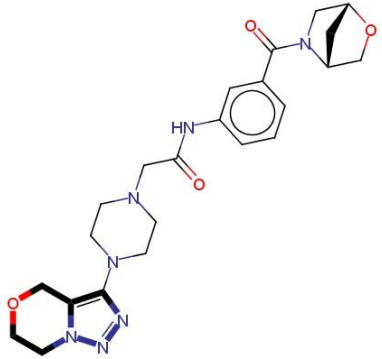
Presence of a 2-(1,2-oxazol-5-yl)pyridine moiety which appears 13 times in the initial dataset



Borderline* 10/11

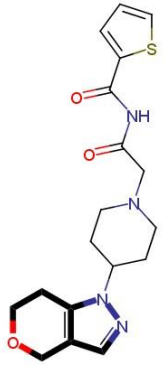


Inactive 8/11



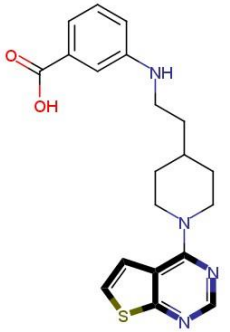
Active 9/11

Presence of a [1,2,3]triazolo[4,3-c][1,4]oxazine moiety which appears 4 times in the initial dataset



Active 8/11

Presence of a pyrano[4,3-c]pyrazole moiety which appears 5 times in the initial dataset

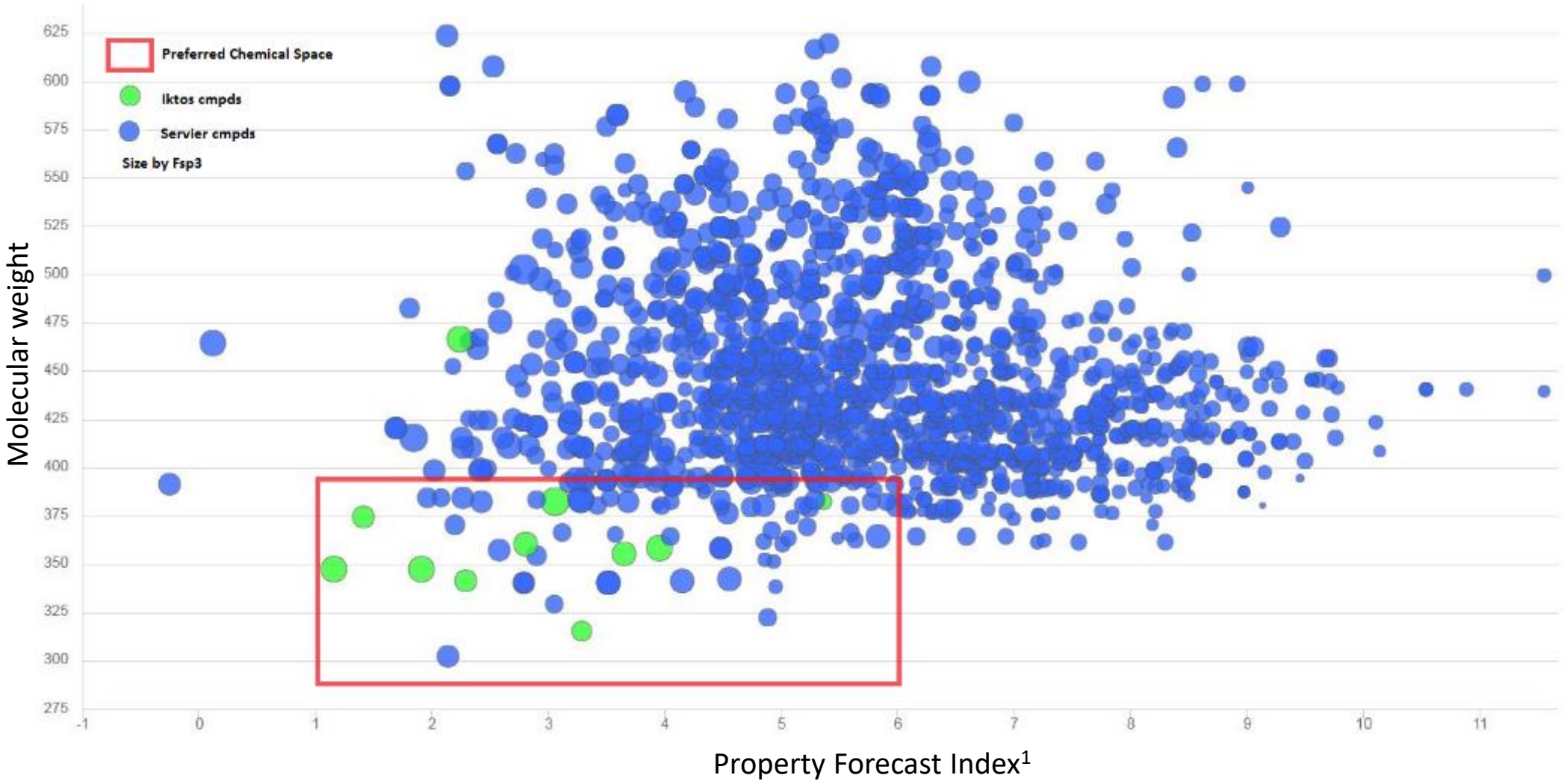


Inactive 9/11

Presence of a thieno[2,3-d]pyrimidine moiety which was absent from the initial dataset

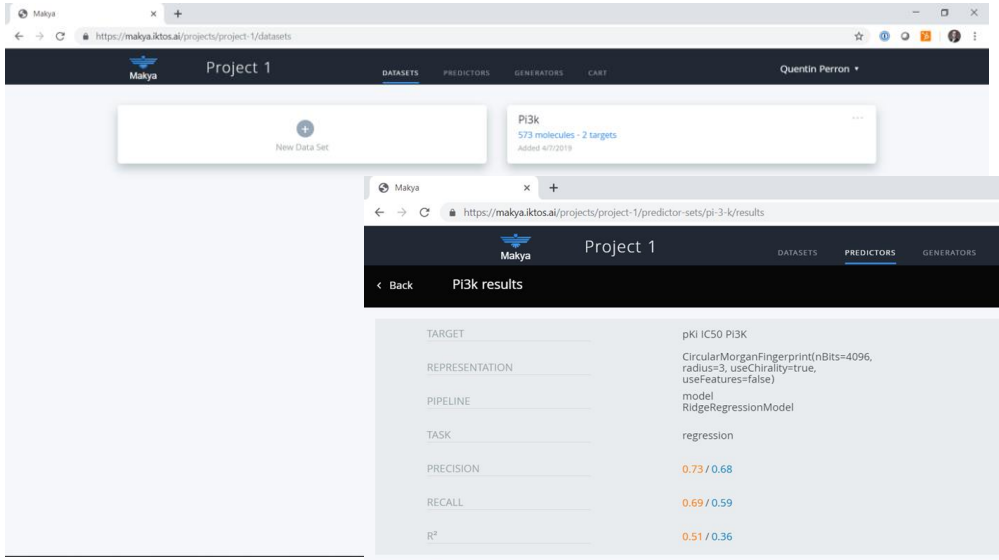
* In the margin error of the Activity assay

Drug-lik

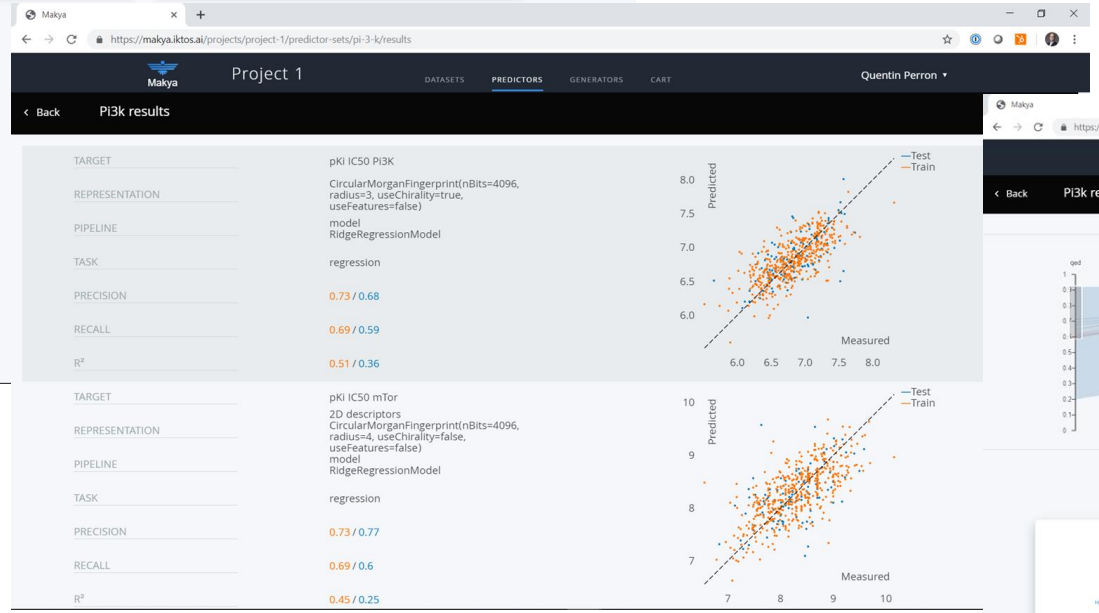


1 - Paul D. Leeson and Robert J. Young Molecular Property Design: Does Everyone Get It? *ACS Med Chem Lett.* **2015,** 6(7), 722–725.

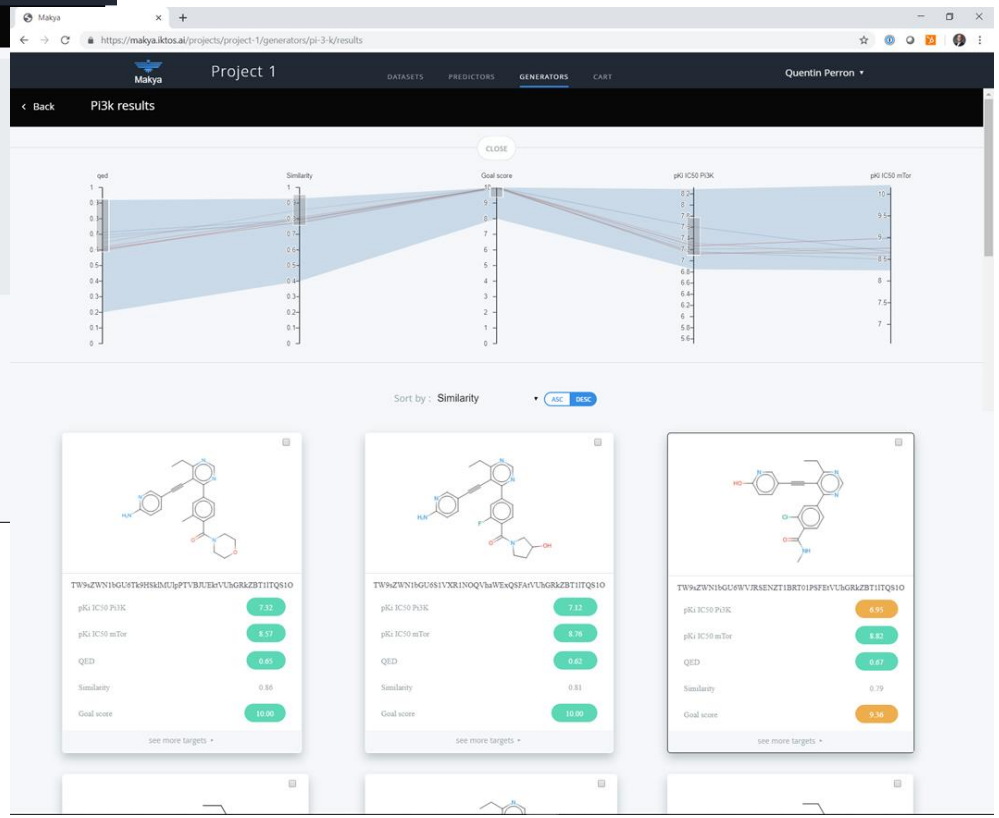
enerative mo



Dataset upload



AutoML module



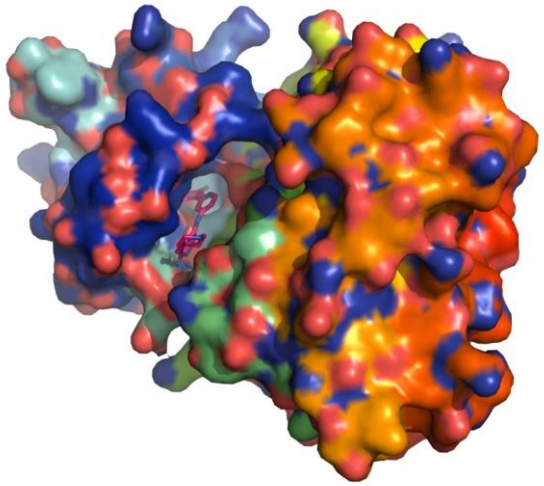
“ideal” in silico propositions

Conclu

- Optimal *in silico* propositions in only a few hours
- An approach which is complementary to the propositions made by the medicinal chemist
- High potential of the technology to accelerate early stage lead optimization – considering several compound design, synthesis and test iterations
- Only local models were used (using the project's data) – likely benefit of integrating in house global predictive models into the process
- Future developments: coupling generative models with structure-based modeling

Perspectives

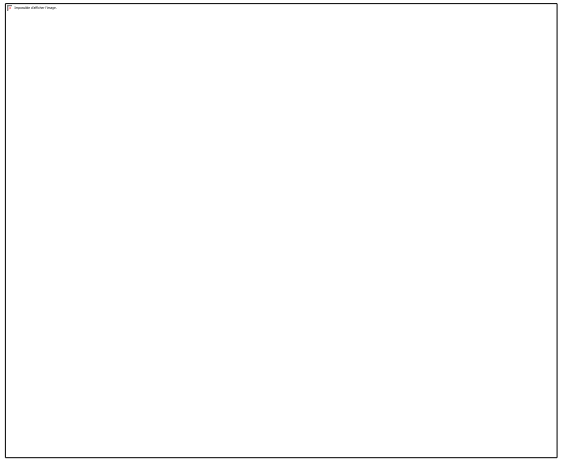
sign with a s



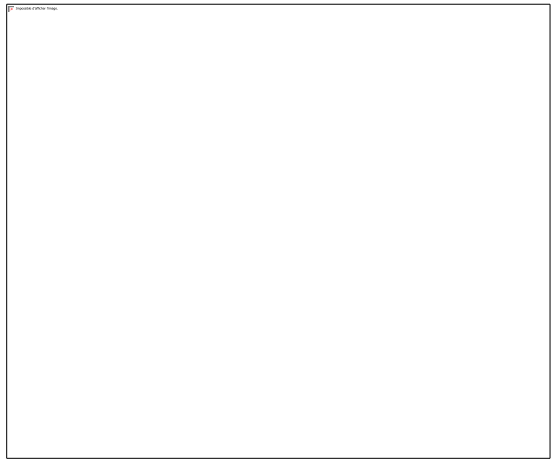
Generator

Docking

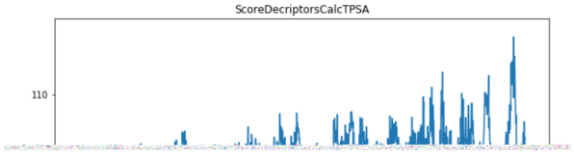
- ✓ High predicted value
- ✓ Important interactions are present
- ✓ All important PhysChem properties are present



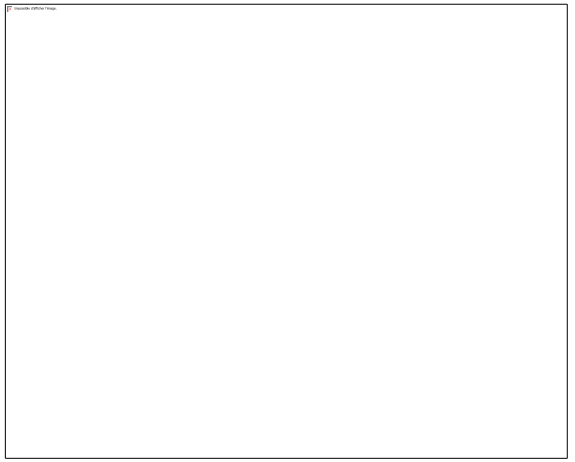
Docking Score



Contact Score

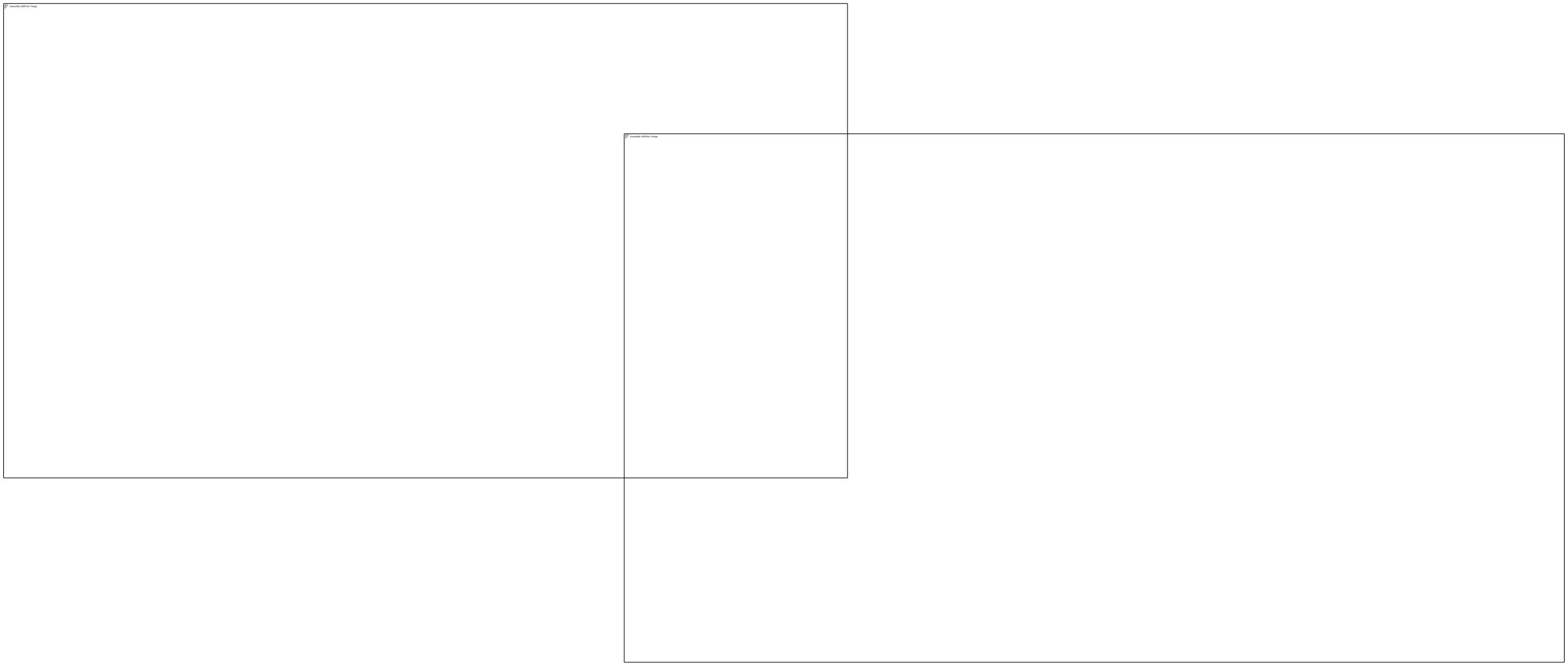


TPSA



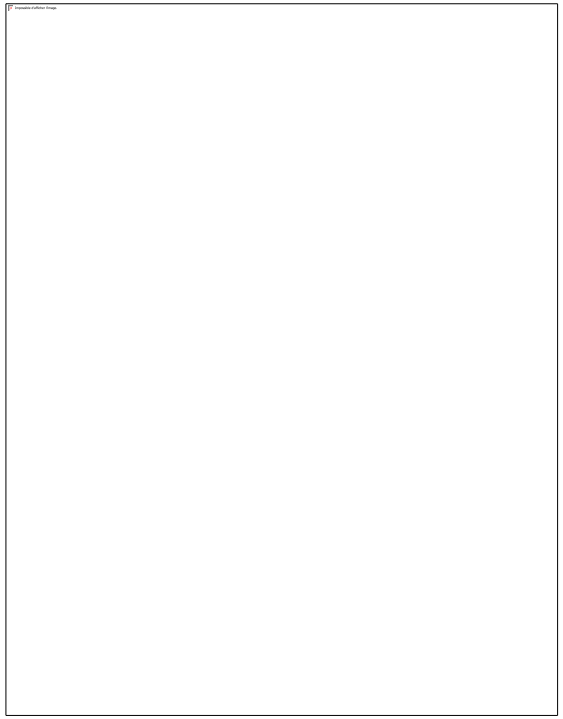
LogP

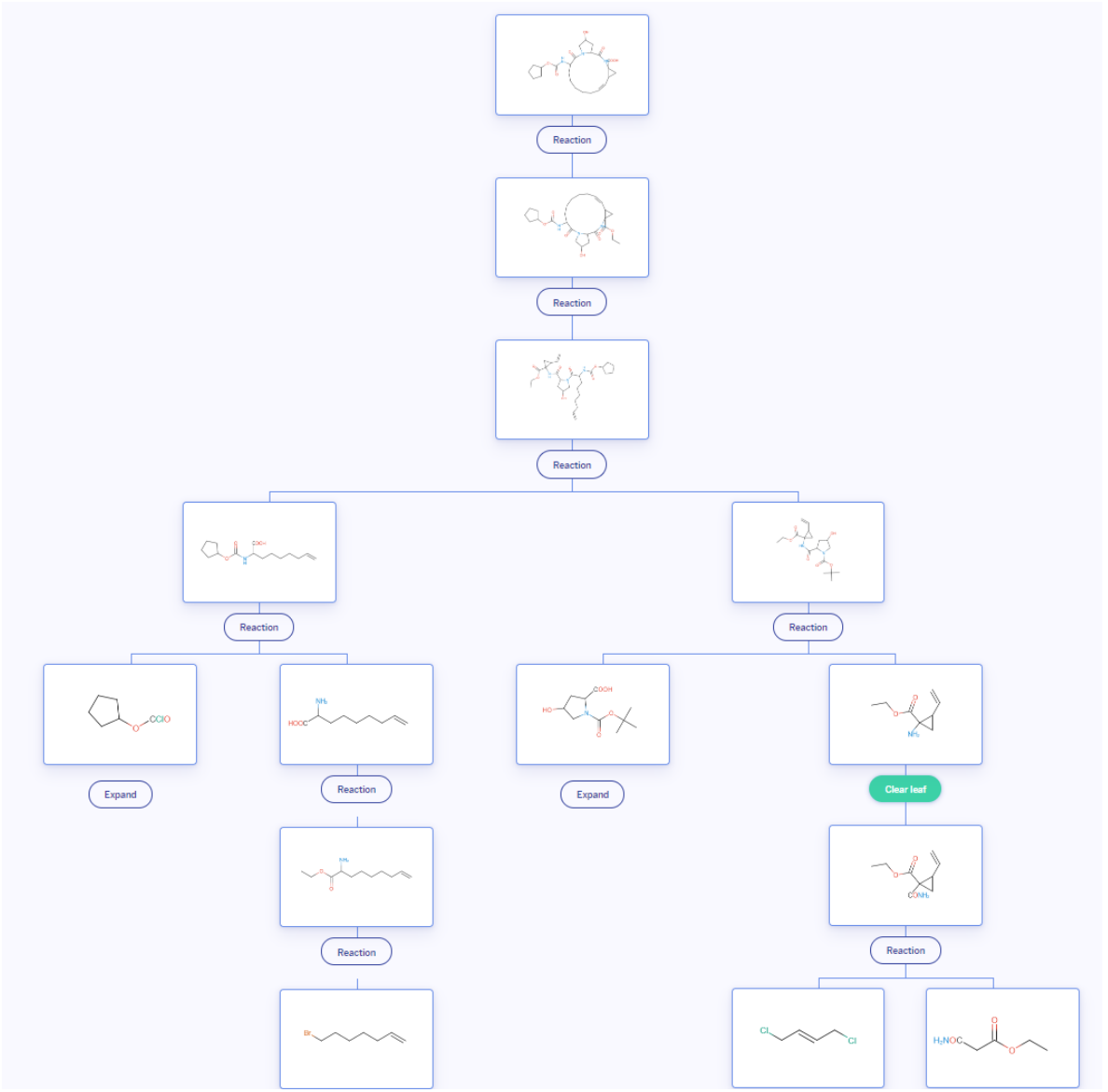
trosynthesis



Freely accessible Beta version coming soon!

Ciluprevir int





Acknowled

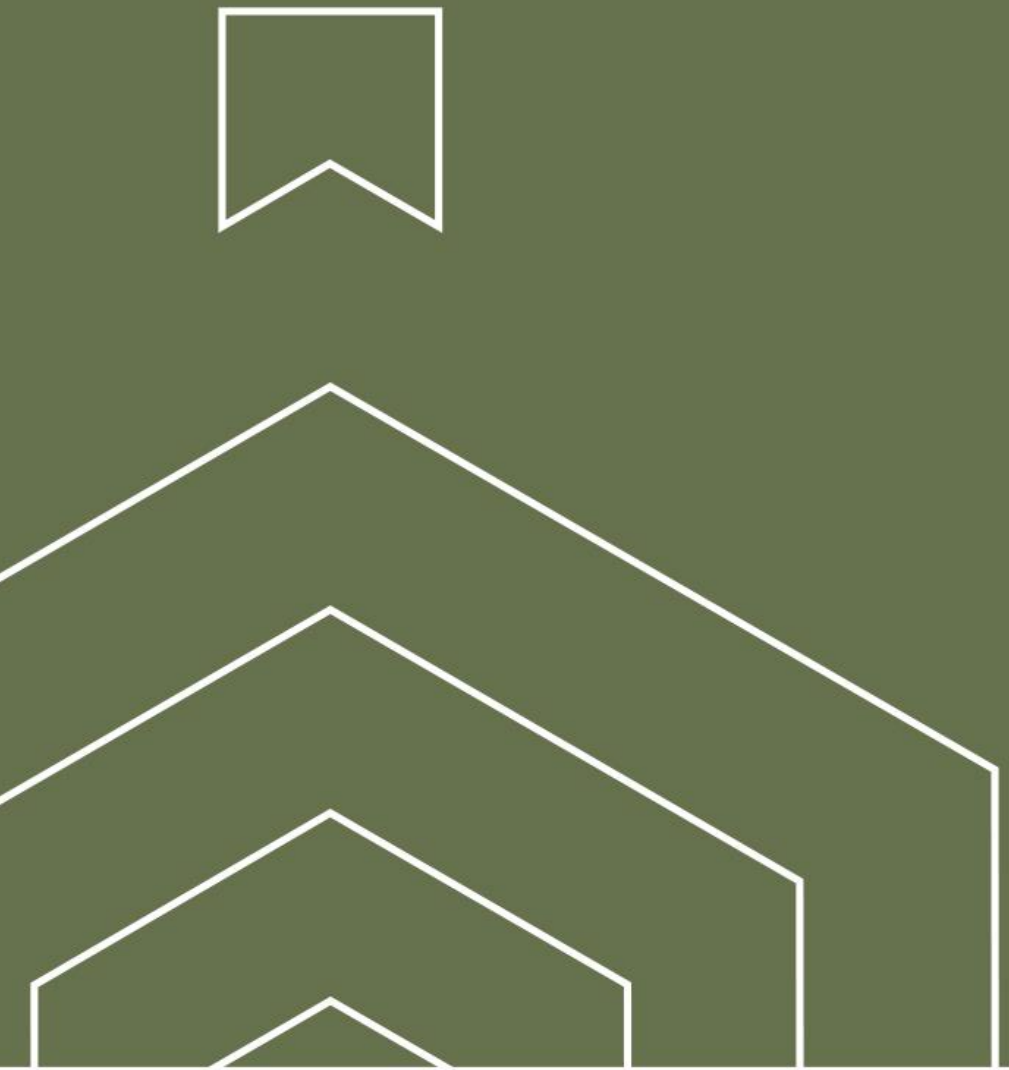
The Iktos team!

Olivier Mirguet, Anne Rojas, Pierre Ducrot, Arnaud Gohier, Marie-Pierre Bourguignon, Patricia Sansilvestri, Philippe Gloanec, Françoise Gellibert

Cont

Quentin Perron
Cofounder & CSO
quentin.perron@iktos.ai
+33 7 68 80 50 76

Back up slides



IKTOS